ABSTRACT OF THE DISCLOSURE

A text classification apparatus directed to a plurality of languages, includes a unit for extracting information for converting a word from non-classified (unlabeled) texts, in a plurality of languages, into a word sense, a unit for learning a classification knowledge at a word sense level after converting a word extracted from a labeled text into a word sense, a unit for learning a classification knowledge at a word level from the labeled text, a unit for learning the classification knowledge at the word level from the classification knowledge at the word sense level and information on a relation between words extracted from the unlabeled text, and a unit for combining the respective classification knowledges to assign a category.